Naval Research Laboratory

Washington, DC 20375-5320

NRL/MR/7220--12-9409

An Alternative Retrieval Algorithm for the Ozone Mapping and Profiler Suite Limb Profiler

Thomas D. Eden
Jerry D. Lumpe

Computational Physics Inc.
Springfield, Virginia

John S. Hornstein

Remote Sensing Physics Branch
Remote Sensing Division
Retired NRL employee

May 1, 2012

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*
01-05-2012

2. REPORT TYPE
Memorandum Report

3. DATES COVERED *(From - To)*
January 2003 – January 2011

4. TITLE AND SUBTITLE
An Alternative Retrieval Algorithm for the Ozone Mapping and Profiler Suite Limb Profiler

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Thomas D. Eden,* Jerry D. Lumpe,* and John S. Hornstein†

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER
72-7105

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Naval Research Laboratory, Code 7227
4555 Overlook Avenue, SW
Washington, DC 20375

8. PERFORMING ORGANIZATION REPORT NUMBER
NRL/MR/7220--12-9409

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Joint Polar Satellite System
NASA GSFC
JPSS/CODE 490.9 Bldg L40/Room #S556
8800 Greenbelt Road
Greenbelt, MD 20771

10. SPONSOR / MONITOR'S ACRONYM(S)
JPSS

11. SPONSOR / MONITOR'S REPORT NUMBER(S)
JPSS 11-065

12. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES
*Computational Physics Inc., 8001 Braddock Road, Springfield, VA 22151-2110
†Retired NRL employee

14. ABSTRACT

Stray light significantly contaminates the signals from many instruments for remote sensing. The usual way of compensating for stray light is to use a model for the production of stray light by the instrument, to estimate what the measured signals would have been in the absence of stray light, and then to use the corrected signals to derive the sought-for geophysical or astrophysical quantities. Correcting the signals for stray light is an ill-posed inverse problem (very close to a deconvolution). In most remote sensing tasks, extracting the geophysical or astrophysical quantities from the signals is also an ill-posed inverse problem. Stacking these two ill-posed inverse problems often leads to unusably unstable results. Stable, accurate results can be obtained by bypassing the estimation of the corrected signals. The geophysical or astrophysical quantities are derived directly from the uncorrected signals, but the production of stray light by the instrument is included in the forward model, along with the external radiative transfer effects that are usually included in the forward mode. That eliminates the stacking of ill-posed inverse problems.

15. SUBJECT TERMS

Remote sensing	Ill-posed inverse problems	Aerosols
Retrieval algorithms	Stratosphere	Stray light
Inverse problems	Ozone	Limb scattering

16. SECURITY CLASSIFICATION OF: a. REPORT	b. ABSTRACT	c. THIS PAGE	17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
Unclassified Unlimited	Unclassified Unlimited	Unclassified Unlimited	Unclassified Unlimited	29	Gerald Nedoluha
					19b. TELEPHONE NUMBER *(include area code)* (202) 767-4246

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std. Z39.18

CONTENTS

AN ALTERNATIVE RETRIEVAL ALGORITHM FOR THE OZONE MAPPING AND PROFILER SUITE LIMB PROFILER

1 INTRODUCTION

The Ozone Mapping Profiler Suite (OMPS) is a coordinated set of instruments that was conceived, designed, built, and tested by Ball Aerospace and Technologies Center (BATC) for the National Polar-orbiting Operational Environmental Satellite System (NPOESS), to ensure that the long-term record of space-based measurements of stratospheric ozone measurements would remain unbroken between the current and next generation of atmospheric sounding missions., and also to provide ozone data for immediate operational uses. OMPS makes three types of measurements: the Ozone Total-column Mapper (nadir viewing, UV wavelengths), the Ozone Profiler (nadir viewing, UV wavelengths), and the finer vertical resolution Limb Profiler (LP: limb viewing, UV, VIS and NIR wavelengths). All three instruments measure ozone absorption features in sunlight scattered by the atmosphere and earth's surface. The first OMPS is scheduled to fly on the NPOESS Preparatory Project (NPP) mission, which is currently scheduled for launch in late 2011.

The OMPS/LP sensor is a six-slit prism spectrometer. Light entering each slit is dispersed by a prism and focused onto a portion of a single, two-dimensional (740×340) charged-coupled device (CCD) focal-plane array, covering a spectral range from 290 to 1000 nm and a nominal altitude range of 0 to 65 km. Light to each slit originates from three separate vertical fields of view (VFOVs) of the atmosphere: one looking backward along the ground track (center slit), and two on either side, separated by 250 km in the cross-track direction. Multiple VFOVs are used to shorten the time between successive measurements of the atmosphere at any given latitude and longitude. Each VFOV is viewed by two slits, one wide and one narrow, to better span the large dynamic range (approximately five orders of magnitude) of the limb-scattered radiance. Thus six images are produced on the OMPS/LP CCD (two from each slit). Both a short and a long integration time are used, to further span the wide dynamic range of the radiance.

Pre-launch characterization of the LP sensor at BATC indicates that off-axis scattered radiation will be a significant problem with the instrument. A general descriptive term for this type of radiation is "stray light." Stray light is defined as scattered radiation, within the instrument, causing photons that emanate from one region of spectral/spatial space to be detected as if they came from another point in spectral/spatial space. In terms of spectral and spatial responses, stray light appears in the wings of a detector's response function. It can be a coherent bump (strong source) or a smooth enhancement of the response wings.

As will be shown, stray light originating from bright regions of the LP CCD (corresponding to wavelength/altitude regions of the atmosphere with high radiance) can easily contaminate low-signal regions, with the potential to affect the quality of the data products. Stray light contamination adversely affects most pixels and is scene-dependent. It must be adequately accounted for in the data processing if the sensor is to achieve the design accuracy for the ozone environmental data record (EDR).

In the standard data processing approach stray light mitigation occurs upstream of the geophysical retrieval algorithm proper, where raw detector signals are converted from counts to radiance units (typically in level 1 processing). This approach usually involves applying a deconvolution algorithm to remove the stray light component from the measured signal, resulting in a corrected signal which is used as input to the level 2 (EDR) retrieval algorithm. The process of deconvolution is well known to be a difficult numerical problem because it is ill-posed, and prone to noise amplification. In low signal conditions, where the stray light could easily be a large fraction of the total signal, errors introduced by this process can often overwhelm the desired information content coming from the primary constituent

Manuscript approved January 28, 2012.

signal of interest. Furthermore, the chain of two ill-posed problems (deconvolution followed by geophysical inversion) is especially vulnerable to noise.

This report documents an alternative retrieval approach for the OMPS/LP data that eliminates the problems associated with signal deconvolution. In this approach all known instrument effects, including stray light, are included in an appropriate forward model to fit the raw (uncorrected) pixel signals directly in the geophysical retrieval algorithm. Unlike traditional retrieval approaches, which use calibrated radiance data as the measurement basis space, this approach utilizes raw (uncalibrated) CCD signals as the measurement basis space. The key difference here is that stray light is modeled in the forward direction to fit the raw detector measurements directly, rather than being removed from the data via deconvolution. We will show that this approach is inherently more stable numerically and generally yields more robust and accurate results than the traditional approach.

To implement this approach, a two-step forward model is needed to generate realistic model detector signals to drive an optimal-estimation retrieval algorithm. The first stage is a radiative-transfer (RT) model that produces idealized radiances, in appropriate spectral and spatial dimensions, from a given atmospheric state and measurement geometry. The second stage is a realistic OMPS/LP instrument model, which propagates this radiance field through the instrument, applying all known instrument optical and electronic effects to produce modeled detector signals, which are then directly comparable to the measured detector data.

Section 2 of this report will outline the fundamental OMPS/LP retrieval problem. The problem will be expressed in the context of a conventional radiance-based retrieval (RBR) scheme. The RBR algorithm was the first algorithm developed by CPI as an alternative to the operational NPP EDR algorithm. In its overall philosophy and approach it is a “traditional” EDR/Level 2 geophysical retrieval, in that it assumes as input a geolocated, calibrated radiance field on a 2D Cartesian grid of wavelength vs. altitude. In this approach the Level 1 algorithms are required to correct for instrument effects such as CCD spectral smile, by interpolation along CCD rows and columns to produce the required Cartesian grid. This algorithm has reached a high level of maturity and provides a valuable testing ground for new ideas, which are then ultimately incorporated into the alternative retrieval algorithm. Section 3 will be devoted to the alternative algorithm, which we call the pixel-based retrieval (PBR) for obvious reasons. Because this retrieval scheme is essentially functionally identical to the RBR algorithm (except for inclusion of the instrument model), the emphasis of this section will be to identify and detail modifications needed to carry out the full retrieval in pixel space on an irregular spatial (altitude) and spectral grid. Retrieval results will be presented in Section 3. Conclusions and future work will be covered in the final section.

2 THE OMPS/LP RETRIEVAL PROBLEM

The purpose of the OMPS/LP sensor is to measure altitude profiles of ozone from the tropopause to the lower mesosphere. The limb scattering technique provides both good vertical resolution and complete global coverage on the day side of the orbit. The first use of this technique was by the Solar Mesospheric Explorer (SME) experiments [1]. Since then the technique has been used by the two Shuttle Ozone Limb Sounding Experiment/Limb Ozone Retrieval Experiments (SOLSE/LORE) [2], the OSIRIS instrument on ODIN [3], by SCIAMACHY [4], and as an experimental part-time mode on SAGE III [5].

The spectral bandwidth of the OMPS/LP measurement encompasses two distinct regions of ozone absorption: UV absorption in the Hartley-Huggins Bands, and visible absorption in the Chappuis Bands. Shown in Fig. 1 are single scattering ozone sensitivities, $d\ln(I)/d\ln(O_3)$, where I is the intensity (equivalent to radiance), for three primary wavelengths used to retrieve ozone. Note the altitude

dependence of the sensitivity curves for these wavelengths. Ozone can be accurately retrieved from UV channels (300 and 318 nm) at altitudes between 30 and 60 km. To extend the ozone retrieval down into the upper troposphere/lower stratosphere the 602-nm visible bands must be included. The OMPS algorithms utilize a number of wavelengths across the instrument spectral bandpass to maximize ozone sensitivity throughout the entire altitude range. As will be discussed below, each spectral region has its own set of difficulties that must be overcome when retrieving ozone. This section will cover these issues in detail, but first we place our algorithm work in its historical context, and then give a top-level description of it.

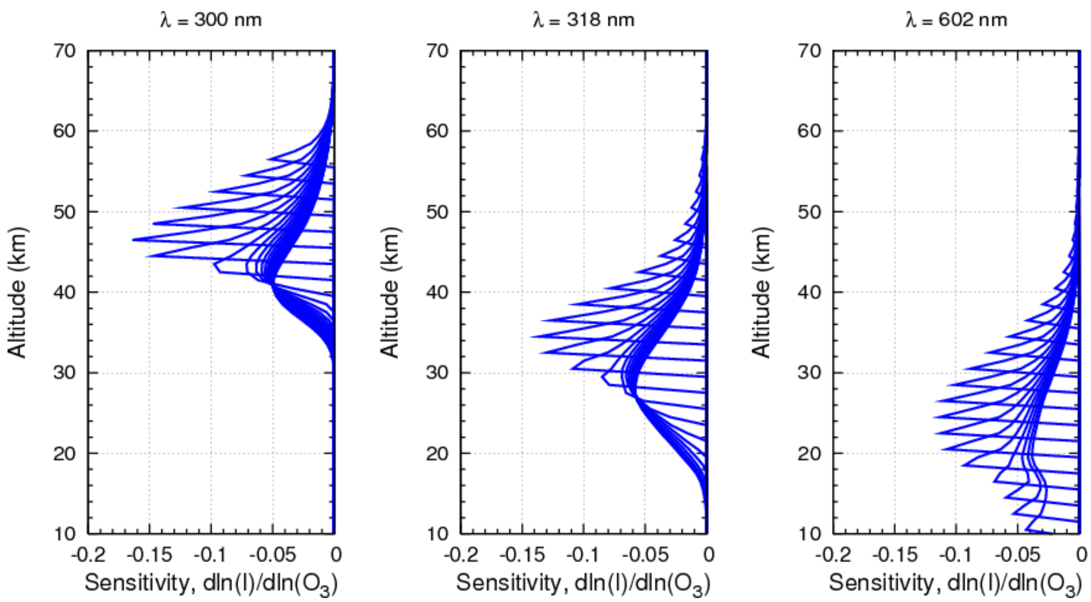

Figure. 1 -- The fractional sensitivity of the change in single-scattering limb radiance, *I*, due to a change in the concentration of ozone for a mid-latitude ozone profile. Each curve is for a different tangent height at a 2-km spacing.

The initial algorithm for the OMPS LP was developed by Raytheon, under contract to BATC [6]. That algorithm was based in part on the algorithm developed for the two SOLSE/LORE flights [2], and included the use of spectral pairs and triplets that had been developed for TOMS and SBUV/2. An alternative algorithm for the OMPS LP was later developed and used in extensive sensitivity studies [5]. These algorithms were all <u>physical retrieval</u> algorithms. A physical retrieval algorithm derives the geophysical profile based on a "best" fit between the measured data and a forward model estimate from a radiative transfer code driven by the distribution of atmospheric composition, surface conditions, and the geometry of solar illumination and satellite viewing. Often the statistics of the scene and of the instrument noise also provide some of the parameters.

Earlier in the OMPS/LP algorithm work CPI developed a <u>regression-based</u> retrieval algorithm [7]. This type of algorithm estimates the profile from the data via a multivariate formula, whose arguments are either the radiances or the signals from the pixels. The parameters in the formula were obtained by adjusting them to give optimal results for a training set of data: a large set of scenes for which the correct profile is known. That is, the formula is obtained by fitting to a training set. No explicit physical model is used; instead, the physics is embodied implicitly, in the members of the training set. (The training set is

obtained either by running a forward model, or from nature.) The fitting can either be explicit (using ordinary fitting methods) or implicit (neural networks or genetic algorithms). In any case a regression-based retrieval algorithm is much faster computationally to apply than is any physics-based retrieval algorithm, because all of the time consuming work is done beforehand, when training the algorithm.

It was discovered that the precision of the algorithm deteriorated dramatically as soon as stray light was included, even if the radiances had been approximately corrected for stray light. This instability results from the stacking of ill-posed inverse calculations; that is, from using the output of one ill-posed inverse calculation (correcting the measured radiances for stray light) as the input for a second ill-posed calculation (calculating the scene from the radiances). To facilitate this work a stripped down mathematical model of the instrument was developed to enable more detailed experiments with algorithms for compensating for the effects of the stray light. In 2006 initial efforts were made to use the instrument model to include stray light in the forward model that generated the training set for the regression-based retrieval algorithm [7]. The expectation was that this would lead to more stable results than trying to correct for stray light an ill-posed inverse calculation of corrected radiances. Results verified that this was indeed the case for the regression-based retrieval algorithm then being tested for the OMPS LP.

It was also noted that no adequate training set yet existed for limb scattering: the representativeness of even a large set of calculated radiances could not be assured, because of the combined variability of the ozone profiles and the aerosol+cloud profiles. Indeed, a representative training set of limb scatter data will only exist after there is a well-validated globally and seasonally representative set of limb scatter measurements that have been taken by an instrument (such as the OMPS LP on NPP) that uses similar wavelengths as the target instrument. Until then, a regression algorithm could not be adequately be trained and tested before launch, and could be trained after launch only if there were an extremely comprehensive (and expensive) campaign of post-launch comparisons with profiles from trusted instruments.

For a regression algorithm, the fundamental idea of including the stray light in the well-posed forward calculation amounts to including the instrumental stray light model in the forward model that is used for calculating the sets of detector signals that constitute the training set. For a physical retrieval algorithm, the instrumental stray light model would instead be included in the forward model that is used in the retrieval algorithm. So this new approach to compensating for stray light could also be implemented in a physical retrieval algorithm. Using experience gained with the regression approach, it was concluded that, at present, a stable physical retrieval algorithm would be both feasible and necessary, despite its computational slowness compared to a regression algorithm. The PBR algorithm that is described in part of this document represents the continued development of those ideas. The initial stripped down instrument model developed for the regression analysis studies became the basis for the detailed instrument model that is now standard throughout the OMPS LP project. The efficiency of this model has been significantly improved through collaborations with other members of the OMPS/LP NPP science team, particularly at Langley Research Center.

2.1 Top-level Algorithm Logic Flow

Shown in Fig. 2 is a top-level logic flow chart of the pixel-based retrieval (PBR) algorithm. Note: If the purple box were eliminated in the optimal-estimation loop, the logic is equivalent to a standard radiance-based retrieval algorithm. The upper path leading to the retrieval algorithm (yellow box) represents the logic necessary to produce simulated data, which would be a radiance field for the RBR algorithm. If the data were intended for the PBR algorithm, a CCD pixel map would be produced. If

realistic instrument noise is to be applied to the radiances, it is applied in a separate procedure using a zero-mean Gaussian random number generator, with knowledge of SNR values acquired during prelaunch calibration. If a noisy CCD pixel map is desired, instrument noise is applied within the instrument model.

The retrieval algorithm is based on an *a posteriori* Bayesian-based inversion theory developed by Rodgers [8]. A modular set of optimal-estimation (OE) routines have been designed and implemented at Computational Physics, Inc. (CPI). These routines have been used extensively to produce operational data products from a number of past remote sensing missions [9-11]. To be compatible with production versions of the OMPS/LP RT and instrument model algorithms, these OE routines have been translated into Fortran 90.

Obtaining accurate ozone retrievals from the OMPS/LP requires that additional geophysical and instrument parameters also be retrieved self-consistently from the measurements. The OMPS/LP retrieval algorithm is a multi-step process in which the following quantities are retrieved sequentially, and in this order (each of these retrieval steps are described in detail in section 2.4):

1. Altitude offset
2. Surface reflectance
3. Aerosol extinction
4. Ozone: Chappuis bands and Hartley-Huggins bands separately

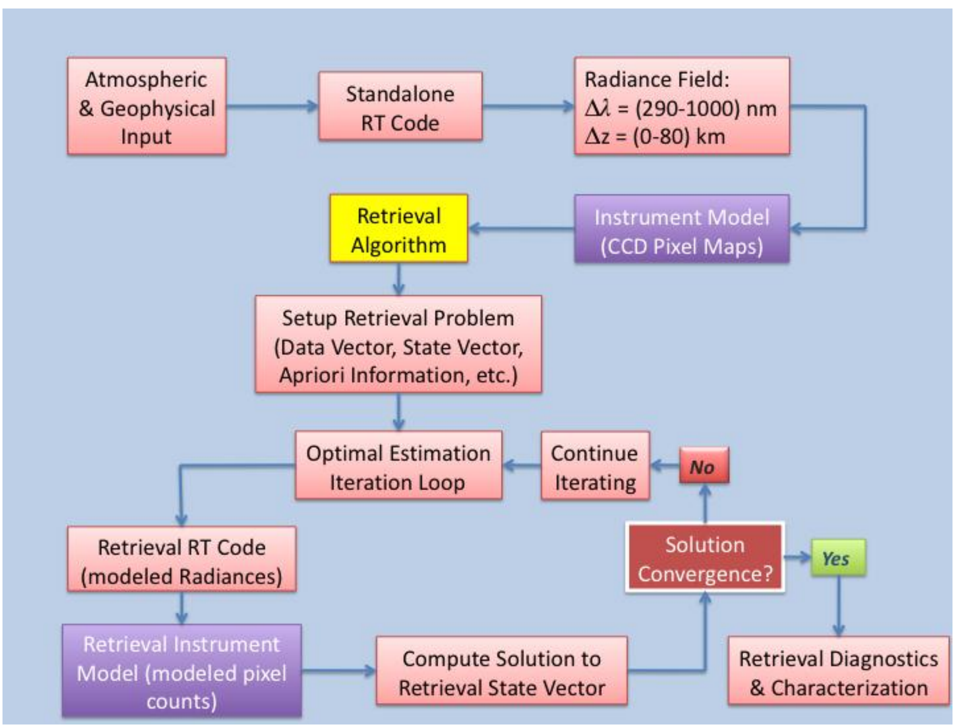

Figure 2 – Top-level flow chart for the PBR algorithm. Eliminating the purple boxes yields the RBR algorithm.

2.2 Radiative Transfer Model

The radiative transfer (RT) model used in these retrieval algorithms is a solar-limb scatter model that includes spherical geometry, all orders of scattering, and polarization [12]. This same model is used as the operational RT code in the OMPS/LP NPP EDR production processing. This model produces simulated radiances, $R_s(\lambda,z)$, as a function of wavelength (λ) and line-of-sight tangent altitude (z) for a given atmospheric state and viewing geometry. The atmosphere is defined by geophysical profiles of temperature, pressure, ozone, aerosols, nitrogen dioxide and other trace gases if desired. The viewing geometry is defined by two fundamental quantities – the solar zenith and azimuth angles. An additional parameter required by the model is the effective surface reflectance. This is an important physical quantity for the solar limb scattering technique, as the visible and near-IR wavelengths have a strong multiple scattering component and solar photons at these wavelengths penetrate to the surface. Thus the measurements are sensitive to the scattering and absorption characteristics of both the atmosphere and the underlying surface scene.

By default the model generates a nominal high-resolution wavelength grid based on two input sources: ozone absorption cross-section data and the solar irradiance spectrum. The grid is constructed to reproduce the highest resolution spectral features in either of these sources. The number of spectral grid points is required to cover each OMPS spectral channel and corresponding bandpass ($\lambda \pm \Delta\lambda$). This grid is used for single-scattering calculations only. The multiple scattering radiance contribution, which is computationally more expensive to evaluate, is generated on a coarser grid and the then interpolated onto the single-scattering grid and added to the single scattering component to yield the total radiance.

Aerosol scattering is included by specifying profiles of aerosol extinction at multiple (minimum of two) wavelengths across the instrument spectral bandwidth. Extinction on the full model spectral grid is then obtained by either interpolation or extrapolation from these input profiles. A simple linear fit to the spectral dependence is then performed, and averaged over the mid-stratosphere to produce a single effective Angstrom coefficient. This is used to infer the mode radius and width of a log-normal size distribution, by interpolation on a look-up table generated by an offline Mie scattering code. These aerosol microphysical properties are used to calculate the aerosol scattering phase function, which governs the aerosol scattering properties in the RT code. Single Rayleigh scattering is also calculated.

Shown in Fig. 3 are radiance spectra at 20 km tangent altitude for a range of solar zenith angles, at two fixed solar-azimuth angles: 0° (left panel) and 40° (right panel).

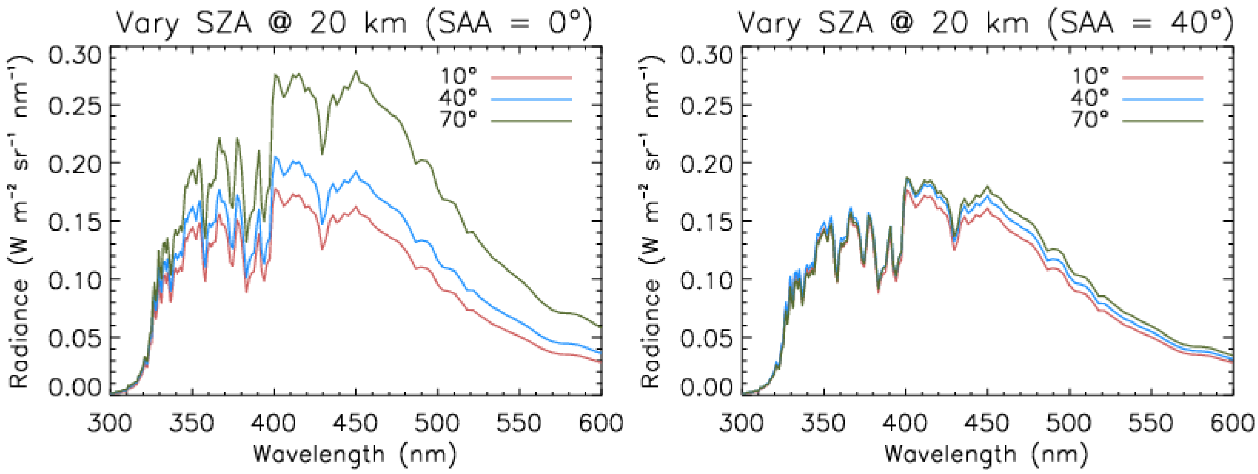

Figure 3 – Radiance spectra produced by the RT code at 20 km for several different solar-zenith angles (SZA), while holding the solar-azimuth angle (SAA) fixed. The calculations used a mid-latitude ozone profile (350 DU) and a moderate springtime aerosol field as inputs.

2.3 Optimal Estimation Package

The RT model is used to drive the optimal estimation module OPT, which performs a regularized inversion to obtain the state vector, *x* that best fits the measurement vector *y*. The relationship between the data and state spaces is defined by the forward model: y=F(x). For most problems of interest this relationship is nonlinear, and the retrieval products are therefore calculated in an iterative procedure. For the i^{th} iteration, the solution can be written as:

$$x_{i+1} = x_a + (S_a^{-1} + K_i^T S_\varepsilon^{-1} K_i)^{-1} K_i^T S_\varepsilon^{-1}[y - F(x_i) + K_i(x_i - x_a)]. \quad (1)$$

Here $y_i = F(x_i)$ is the data estimate calculated by the RT model, K_i is a derivative matrix (kernel) describing the sensitivity of $F(x_i)$ to changes in x_i, K=dy/dx, x_a is the *a priori* value of *x*, with covariance S_a, and S_e is the measurement covariance matrix. The superscript *T* denotes a matrix transpose operation.

The ozone (UV and visible) and aerosol extinction retrievals utilize single-scattering analytic kernels generated by the RT code. The main reason for using these kernels is to increase retrieval speed, as the large number of forward model calls required to construct numerical (finite-difference) kernels can be the primary CPU driver in the data processing. The analytic kernels, since they include only single-scattering, are an approximation to the full scattering problem; however, the retrieval errors caused by this approximation are found to be small. Since analytic kernels are not available for the surface reflectance and altitude offset retrieval problems, the full scattering, finite-difference kernels are generated for these retrievals. This does not seriously impact overall processing time because these problems are much faster due to the simple single-element state vector (see section 2.4).

In the optimal estimation technique the solution vector represents an optimal combination of direct information from the data and the *a priori* estimate of the state. The relative weight given to these two information sources is determined by the data and *a priori* covariance matrices, S_e and S_a. The diagonal elements of S_e are computed from knowledge of the instrument signal-to-noise (SNR) values obtained from prelaunch calibration at BATC. Off-diagonal elements, which represent spatial and spectral correlations in the data, are currently set to zero. The diagonal components of S_a are assigned values by multiplying the i^{th} term of x_a by a relative error factor, δ. In other words, the *a priori* variance is assumed to have a constant fractional value. The off-diagonal elements of S_a, corresponding to altitude coupling in

the retrieved profile, are used only for the aerosol extinction and ozone retrievals and are calculated as follows:

$$S_a^{ij} = [S_a^i S_a^j]^{1/2} \exp\left[-\left|z_i - z_j\right|/l\right] \tag{2}$$

Here, l is the covariance length, which is dependent on the retrieved constituent, and z_i and z_j are altitude levels corresponding to the i^{th} and j^{th} elements of S_a, respectively.

Criteria for convergence depend on many factors and there is no unique way to define a single best convergence criteria for all retrieval problems. The OPT routines check for convergence after every iteration, and for the OMPS/LP retrievals we have implemented a criteria based on the following inequality:

$$K_i \Delta x_i \le \tau [S_\varepsilon^{ii}]^{1/2} \tag{3}$$

Here, $\Delta x_i = x_{i+1} - x_i$, is the difference between the current iterate of the state vector x_{i+1} and the previous iterate x_i, and τ is a user-specified tolerance factor. This relation simply implies that convergence is reached when the incremental change in the forward model data vector is less than a specified fraction of the measurement variance. That fraction, τ, is for the most part empirically determined by exercising the retrieval to best achieve the prescribed accuracy in the fewest iterations.

2.4 OMPS/LP Retrieval Products

The order of the four retrieval steps outlined in Section 2.1 is important, and has been chosen to yield the most robust ozone retrieval results in the minimum amount of steps. For example, for the first step in the process, the altitude offset retrieval, all other quantities to be retrieved are still unknown and therefore by necessity it uses only *a priori* information for the atmospheric state. Subsequent steps, however, must incorporate updates to the instrument/atmosphere state vector obtained from previous retrieval steps. In particular, the Chappuis-band ozone retrieval is strongly affected by the interfering aerosol component, so it is crucial that this step incorporate a good first estimate of the aerosol field obtained from step 3 rather than simply fixing the aerosols to a climatological *a priori* value. A detailed discussion of each separate retrieval problem is given below.

2.4.1 Altitude offset

Unlike some limb-viewing remote sensing instruments, the OMPS/LP does not actively scan the earth's limb but stares at a fixed elevation angle. Projection of the finite vertical field-of-view of the slit onto the limb determines the range of altitudes sampled, and there is a fixed mapping function between elevation offset angle, line-of-sight tangent altitude, and CCD pixel number. Multiple possible error sources - in pre-launch characterization of the relative instrument/satellite alignment, satellite ephemeris, or telemetry timestamps – can cause an error in the absolute altitude registration. Because of the fixed nature of the CCD array, altitude errors for the LP will take the form of an overall altitude shift, characterized by a single parameter Δz.

We use a technique called Raylegh Scattering Altitude Sensor (RSAS) to retrieve this altitude shift directly from the LP data for each orbit. This step is performed first, as the subsequent ozone and aerosol retrievals depend sensitively on having accurate altitude registration. As the name implies, the retrieval is based on fitting a radiance profile that is dominated by the single-scattering Rayleigh signal, i.e., using an altitude and wavelength range where ozone absorption and aerosol scattering are negligible. These requirements are well satisfied by looking at wavelengths near 350 nm and altitudes of 40 km and above.

This wavelength corresponds to hole in the ozone absorption spectrum and at this altitude there is no detectable stratospheric sulfate aerosol. Of course, this approach requires that the atmospheric temperature/pressure profiles are known to sufficient accuracy that the modeled Rayleigh scattered radiance has less error than the altitude-shifted measurement. Since the temperature/pressure is taken from an independent stratospheric analysis (e.g., NCEP or UKMO), there is good reason to push the retrieval to the lowest altitude possible (but still above the aerosol layer) where these analyses are the most accurate. The commonly accepted uncertainties in the analyses are sufficient to allow correction of pointing errors to the level required for the ozone retrieval.

The simplest data vector that can be constructed for this problem is a single radiance profile at one wavelength (λ_o), between 40 and 50 km; namely, $y=R(\lambda_o,z)$. The state vector consists of a single parameter, $x=\Delta z$. The modeled data vector, $y_0=F(x_0)$, is then calculated from a single call to the retrieval RT code using the initial value x=0 (no offset). Because the atmospheric state does not change during the course of the retrieval, no other calls to the RT code are required. Subsequent iterations y_{n+1} are fit to y by simply vertically shifting the model radiance profile by a fixed altitude perturbation and interpolating to the new grid. The kernel matrix K_n is constructed from the following finite-difference ratio:

$$K_n = \frac{\Delta y}{\Delta x} = \frac{(y_n - y_{n-1})}{(x_n - x_{n-1})} \tag{4}$$

Here, Δx is the perturbation of the state vector, which for this case is taken to be a fixed altitude increment. This retrieval is extremely quick and reaches convergence in just a couple of iterations.

2.4.2 Surface Reflectance

The spectral dependence of the underlying surface reflectance is an important parameter for the subsequent ozone and aerosol retrieval problems, and therefore it is important to retrieve it self-consistently from the data rather than fixing it to a climatological value. This is particularly true when the surface scene below the instrument footprint is complex and inhomogeneous, which for example is the case for broken cloud fields and mixed ocean/ice/land terrain. Currently, the RT code assumes a constant surface reflectance over all visible and NIR wavelengths used by the LP, typically from 480 to 1000 nm. This situation is obviously over-simplified and should be modified in the future to incorporate a more realistic spectral behavior of the surface albedo.

For the time being, working with the existing model, we retrieve a single surface reflectance value and therefore the state vector for this problem, like the altitude retrieval problem, is a simple scalar. For the data vector we use the measured radiance at a fixed altitude level of ~40 km in three spectral windows in the visible region:

a. 480-550 nm

b. 630-758 nm

c. 768-830 nm

Sensitivity to surface properties comes from the multiple scattering component of the measured radiance, since photons need to be scattered at least twice to encounter the surface and then be scattered into the instrument FOV, The UV channels are avoided because the lower atmosphere is opaque at these wavelengths, which are completely dominated by single scattering and thus contain no information on surface reflectance. Wavelengths between 550 and 630 nm are avoided because of strong ozone

absorption, and similarly the spectral region between 759 and 767 nm is avoided due to the O_2 *A*-band. The 40-km altitude level is chosen because it is above any significant source of single-scattered opacity (i.e., aerosols) at the visible and near-IR wavelengths.

The data vector in this case consists of a radiance spectrum at one altitude, y=[R(λ_1,40 km), …, R(λ_{NY},40 km)]. The surface reflectance kernel is calculated numerically using the finite difference scheme of Eq. (4). The state vector perturbation is again taken to be a constant value in surface reflectance (0.1), but in this case the RT code is called at each iteration to re-generate the 40-km radiance spectrum.

2.4.3 Aerosol Extinction

Knowledge of the loading and spectral behavior of aerosol extinction from the upper troposphere through the stratosphere is critical for retrieving ozone in this region. Aerosol scattering is a continuum, slowly varying (~1/λ) contribution to the background atmospheric extinction in the stratosphere and thus, like Rayleigh scattering, is present at all wavelengths. It must be characterized and removed from the primary ozone channels or it can significantly bias the ozone retrievals, particularly under conditions of high loading such as after volcanic eruptions or in the presence of polar stratospheric clouds (PSC). Because the aerosol component is highly variable in space and time, it must be retrieved directly from the OMPS data in parallel with the ozone retrieval. This requires the use of window channels in the ozone absorption spectrum because the aerosol extinction, while a significant contaminant, is still dominated by ozone in the primary ozone absorption regions.

The goal of this step is to characterize the aerosol spectral dependence across the entire LP spectral range so that it can be removed from all channels used in the ozone retrievals. At a minimum, two extinction profiles must be retrieved to allow for spectral interpolation to non-aerosol channels, and to compute the effective Angstrom coefficient needed to derive the aerosol microphysical parameters for the aerosol scattering phase function (see Section 2.2). For each wavelength profile, we construct a data vector by normalizing the radiance profile to a reference tangent height to reduce sensitivity from pixel calibration errors and surface reflection:

$$y = \ln\left[\frac{R(\lambda,z)}{R(\lambda,z')}\right] \tag{5}$$

Here, *R* is the radiance value at wavelength λ and altitude *z*, and *z'* is the normalization altitude.

The cleanest aerosol signature in the OMPS/LP spectral bandpass will come from the longest wavelength channel available. Although the detector is nominally sensitive out to 1 μm, the high-wavelength cutoff is effectively at 840 nm. This limitation is due to the reduction in throughput by the IR blocking filter, which is employed to reduce stray light. Therefore, the high-wavelength aerosol channel was taken to be 840 nm. For the low-wavelength channel we have used 455 nm. This avoids the strong ozone absorption in the Chappuis bands but is not too far into the UV, where the aerosol sensitivity decreases rapidly due to the high optical opacity in the stratosphere.

Shown in Fig. 4 are results from a Chappuis-band ozone retrieval using the RBR algorithm. The simulated data used as input were produced from a 150-DU ozone profile, along with associated temperature and pressure profiles, a solar zenith angle of 0.0^o, a solar azimuth angle of 70^o, and a surface reflectance of 0.1. The simulated data included realistic random noise. The error profiles in Fig. 4 illustrated the importance of performing the aerosol retrieval prior to the ozone retrieval. Reversing the order of these two retrieval problems (*i.e.*, retrieve ozone first, then aerosol extinction) does not improve

the outcome. The order of the retrievals matter because aerosol extinction perturbs the ozone retrieval more than ozone absorption perturbs the aerosol retrieval.

The use of a linear parameterization (aerosol extinction retrieval at two wavelengths) to estimate the aerosol field, while perhaps sufficient to effectively correct for aerosols in the ozone retrieval, would not produce a science-quality aerosol extinction EDR product. The spectral behavior of realistic stratospheric aerosol extinction has a more complicated nonlinear signature. The aerosol retrieval module has been tested using a second-order polynomial parameterization for the spectral dependence and this method yields better results overall. However, this approach is currently impractical for operational implementation of the algorithm since single scattering analytic kernels need to be utilized to increase retrieval speed. Coupling between the three retrieved profiles in the quadratic parameterization is fairly strong, resulting in large off-diagonal components to the kernel matrix. The current operational RT model does not produce any analytical coupling of the kernels between aerosol extinction profiles. This coupling needs to be coded in explicitly in terms of the three single-channel analytic kernels. At the time of this writing, this approach has not been implemented successfully. There is no problem if the kernel is computed numerically with the finite-difference scheme. Further work is required to bring this type of retrieval into production mode.

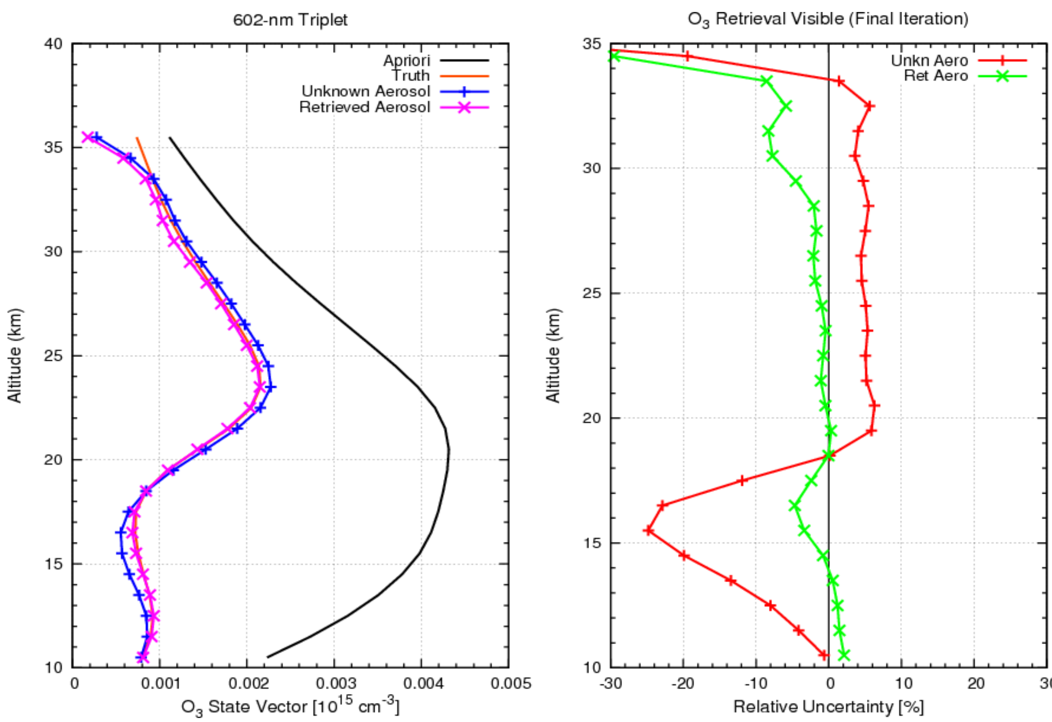

Figure. 4 — Visible ozone retrieval results using the RBR algorithm. As defined, "unknown aerosol" denotes an unknown aerosol *a priori* used in the ozone retrieval, and "retrieved aerosol" is defined as aerosol that was retrieved (using the unknown aerosol *a priori* and unknown ozone), and used as input to the ozone retrieval. The results shown in the right-hand panel indicate the need to retrieve aerosol extinction prior to retrieving ozone in the OMPS/LP retrieval chain.

2.4.4 Chappuis-band Ozone Retrievals

The primary goal of this retrieval problem is to maximize the algorithm's sensitivity to the underlying ozone signature by reducing or eliminating the effect of competing absorption/scattering processes in this spectral region. The dominant interference terms are extinction due to aerosol and Rayleigh scattering. Both of these processes are characterized by a slowing varying spectral signature, in

contrast to the more structured ozone absorption spectrum. An effective way to enhance the ozone signature at the expense of the background aerosol and Rayleigh components is to construct the data vector *y* from a "triplet" of channels, using tangent-height normalized radiance profiles:

$$y = \ln\left[\frac{R(\lambda_1,z)}{R(\lambda_1,z')}\right] - \ln\left\{\left[\frac{R(\lambda_L,z)}{R(\lambda_L,z')}\right]^{\omega_L} \cdot \left[\frac{R(\lambda_R,z)}{R(\lambda_R,z')}\right]^{\omega_R}\right\}. \tag{6}$$

Here the notation has the same meaning as in Eq. (5). The wavelengths employed are: 1) high ozone absorption, λ_1 = 602 nm; 2) low ozone absorption (short wavelength), λ_L = 499 nm; and 3) low ozone absorption (long wavelength); λ_R = 675 nm. The use of closely spaced high- and low-absorption ozone channels helps maximize the differential ozone absorption while canceling the relatively slow-varying aerosol term. The weighting factors w_L and w_R are given by:

$$\omega_L = \frac{\lambda_R - \lambda_1}{\lambda_R - \lambda_L} \text{ and } \omega_R = 1 - \omega_L \tag{7}$$

The altitude range for this retrieval is from 10-35 km. Single-scattering analytic kernels from the RT model are appropriately transformed to be used in the optimal estimation component of the retrieval algorithm.

2.4.6 Hartley-Huggins Band Ozone Retrievals

Rayleigh scattering is the dominant competing effect when retrieving ozone in the UV Hartley-Huggins bands. In reducing the effects of Rayleigh scattering, a tangent-height normalized "pair" wavelength grouping is employed to construct the data vector *y*:

$$y = \ln\left[\frac{R(\lambda,z)}{R(\lambda,z')}\right] - \ln\left[\frac{R(\lambda_0,z)}{R(\lambda_0,z')}\right]. \tag{8}$$

Except for λ_0, all other variables are identical to those used in the triplet data vector in Eq. 6. The primary "on" channel λ corresponds to a high ozone absorption coefficient, while the reference channel λ_0 (=350 nm), is used as a low-absorption contrasting channel. This is a very simple differential measurement, with the background Rayleigh scattering terms largely canceling between the two channels.

Because the ozone absorption cross section changes rapidly as a function of wavelength throughout the UV bands, and the altitude sensitivity of the retrievals varies accordingly (see Figure 1), four different wavelengths are used to maximize the retrieval sensitivity over the altitude range from 30 to 60 km. The wavelengths currently used include: 294, 300, 310 and 320 nm.

2.5 Simulated Data Ensemble

Simulated data are needed for exercising and characterizing the retrieval algorithm. With simulated data, the retrieval results can be compared directly to the known ("true") atmospheric state, allowing for an unambiguous characterization of retrieval error. Referring back to Fig. 2, the production of simulated data follows the path leading up to the *Retrieval Algorithm* box. The simulated data mode utilizes a standalone version of the RT code for producing simulated radiances. If simulated CCD maps are needed, then an instrument model is used as a second step (see Section 3).

Constructing a simulated data ensemble with diverse atmospheric and viewing geometry inputs provides a realistic testing ground for the retrieval algorithms. The heart of the ensemble used here consists of three different ozone profiles: 1) a low-latitude 550-DU profile, 2) a mid-latitude 350-DU profile, and 3) a high-latitude 150-DU profile. Each of these profiles has associated profiles of atmospheric pressure and temperature. Fig. 5 shows an example of how the ozone, pressure, and temperature profiles were created by averaging the 325 and 375 DU reference profiles in the OMPS climatology. For each set of ozone, pressure, and temperature profiles, 72 simulated data runs were generated from different combinations of atmospheric state and viewing geometry, leading to a total of 216 runs for the full ensemble. Listed below are the parameters that were used in different combinations with each ozone profile.

- Aerosols: Two sets of seasonally averaged POAM aerosol data (spring and winter) were used in constructing the ensemble data sets. Three data sets for each seasonal period were generated from three different mode radii (0.07, 0.10, and 0.16 μm) for the stratospheric aerosol particles using a POAM 600-nm reference profile in a Mie scattering calculation. Shown in Fig. 6, are the seasonal spectral profiles of aerosol extinction at 20 km generated from the POAM data.
- Surface reflectance: 0.10 and 0.70
- Solar zenith angle: 10°, 40°, and 70°
- Solar azimuth angle: 0° and 40°

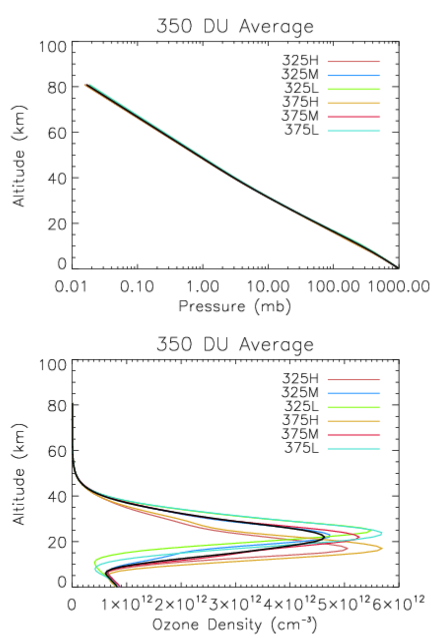

Figure 5 – Average pressure (upper left panel), temperature (upper right panel), and ozone (lower left panel) profiles. Solid black lines denote the averaged profiles from the six individual profiles shown.

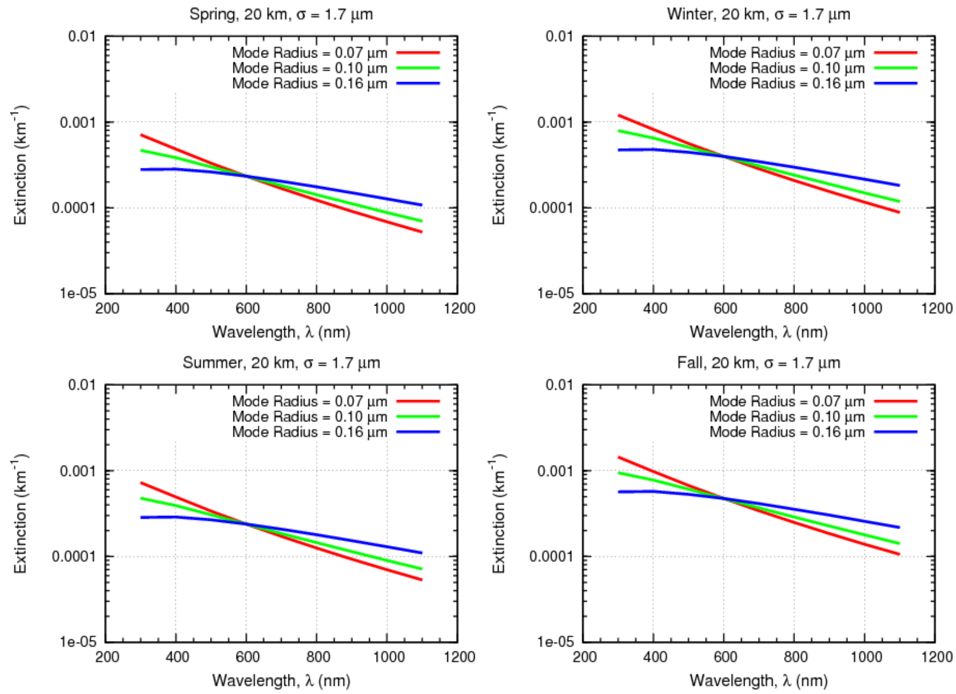

Figure 6 – Modeled POAM aerosol extinction spectral distributions at 20 km. The spring and winter seasons were used as inputs for generating ensemble data.

Figure 7 shows results of simulated ozone retrievals using this 72-run data ensemble. The upper panel shows the Chappuis-band ozone retrieval using a mid-latitude 350-DU ozone profile as “truth”, with realistic random noise applied to the model data. The bottom panels show the results of the UV ozone retrieval from the same set of simulations. A retrieved surface reflectance was used as input to these retrievals, and the input aerosol field was identical to what was used to create the simulated data.

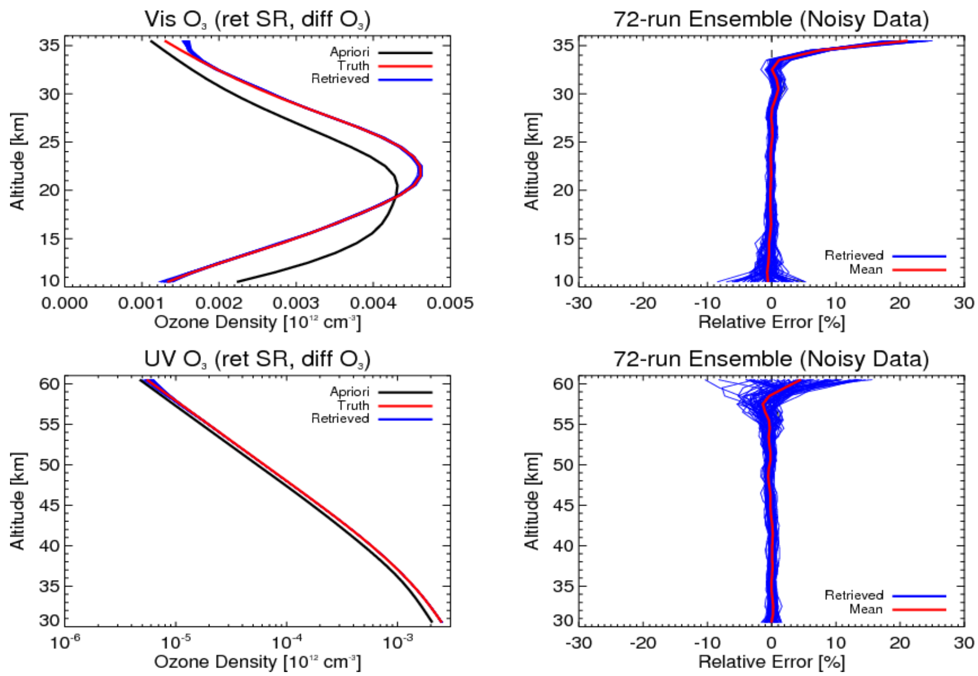

Figure 7 — RBR algorithm ozone retrieval results, using the 72-run ensemble data set with noise applied to each data set. A retrieved surface reflectance and different ozone *a priori* profile were used as inputs. All other relevant inputs were fixed at "truth" values

3 PIXEL-BASED RETRIEVAL ALGORITHM

The RBR algorithm is designed to process radiance data that is noisy but otherwise correct (i.e., contains no systematic errors). When the radiance data are contaminated by stray light, and that stray light cannot be adequately corrected for, the RBR algorithm will give incorrect and unstable results. This has been confirmed by experiments in which data contaminated by stray light was fed to the RBR algorithm. And, of course, this agrees with the conclusion drawn from earlier experiments using the regression-based algorithm.

The effect of stray light is important in limb-scatter measurements because of the large dynamic range spanned by the true radiances over the relevant range of tangent altitudes and wavelengths. Because of the large dynamic range, relatively small amounts of stray light in pixels with low signal can have large negative effects. Furthermore, because of the three-VFOV design, stray light can cause light that should have fallen onto the bright portion of the spectral image from one VFOV to fall instead on the dim portion of the spectral image from another VFOV.

As described previously, we have developed an alternative approach in which stray light is applied in the retrieval forward model, *i.e.*, it is convolved with the CCD field, which is a well-behaved mathematical operation. The critical element in this alternative PBR algorithm is the OMPS/LP instrument model, which is described in detail in Appendix A. Stray light is characterized by the point-spread functions (PSF) of the optical system and the instrument model incorporates the full set of 108 PSFs measured by BATC during prelaunch calibration. The primary goal in testing the PBR algorithm has been to retrieve ozone. Using the ensemble of simulated data described in Section 2.5, with the additional step of producing simulated CCD maps from the radiance data with stray light applied, much has been learned about the strengths and challenges in the PBR approach to the LP EDR algorithm.

All results shown in this section used a retrieval forward model that has stray light convolution turned either on or off. These tests give an idea of just how much stray light convolution matters to retrieval accuracy. A call to the two-step retrieval forward model is fairly time consuming because all pixels, which are essentially at different wavelengths, must be calculated. The RT code must be called first to calculate radiance profiles for all wavelengths in the spectral bandwidth of the instrument. Second, a full PSF convolution is performed in the instrument model, which is also CPU intensive.

Shown in Fig. 8 is a contour plot showing the relative fraction (in percent) of stray light contaminating the pixels of the OMPS/LP CCD for a typical atmospheric scene. Essentially, it represents the relative difference of CCD pixel signals in two runs with and without the application of stray light. Note that each slit is seeing the same atmospheric view for this analysis (i.e., there is no effort to simulate a different atmosphere in the three slits). All pixels with stray light contribution greater than 100% are put in the 100% bin; hence, the red saturated areas that appear in each image. The important point here is the amount of stray light contamination contained in the white rectangles that appear in the central images. The two rectangles in the bottom central image represent the pixels used to retrieve ozone. The rectangle on the left(right) is where the Chappuis(UV)-band ozone is to be retrieved. In each rectangle, stray light contamination on the order of 20% is clearly evident. For UV ozone, low SNR values for many of these pixels results in complications that make a retrieval of ozone nearly impossible to obtain.

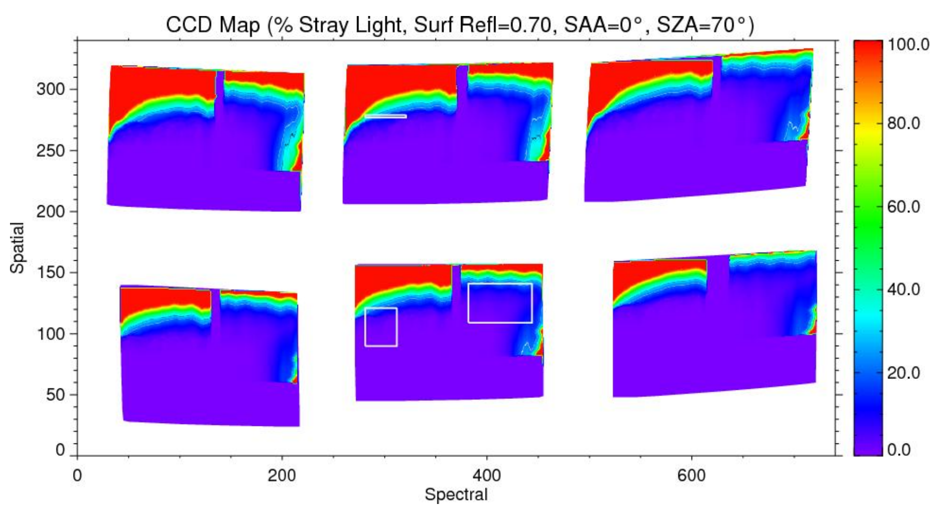

Figure 8 – CCD signal map showing percent stray light for the six images on the OMPS focal plane. This map was produced using the 150-DU ozone profile with a springtime aerosol field, and other input parameters that are listed at the top of the plot. White boxes in the central images indicate pixels that are used in the retrievals of ozone and surface reflectance. The left-hand box in the lower central image denotes the area from which Chappuis-band ozone is retrieved; the right-hand box in the same image denotes the area from which UV-band ozone is retrieved. The small thin white box in the upper central image denotes pixels that are used in the retrieval of surface reflectance.

3.1 Significant Modifications to PBR Formalism

Much of the logic (*e.g.*, optimal estimation) of the PBR algorithm is identical to that of the RBR algorithm. The same four retrieval problems are executed in the same order. However, apart from the instrument model, the biggest changes in logic are observed in the calculation of the data vectors. There are two reasons for these changes. First, the raw CCD images naturally produce an irregular pixel grid in spatial/spectral space (spectral "smile" etc); and second, because of low SNR values for high-altitude

pixels used for normalization, a certain amount of averaging is necessary to obtain stable high-altitude normalization factors.

Another fairly significant modification was made to the single-scattering analytic kernels used in the optimal estimation. This change was required because, in addition to including the production of stray light, we chose to use the uncalibrated signals instead of the calibrated radiances, and also to include the effects of spectral and spatial smile in the instrument model. Because of the former decision, the values of the kernel must be appropriate to signals in counts rather than to calibrated radiances and irradiances. Because of the latter decision, in the PBR algorithm instrument model the independent variables of the kernels must be the row and column indices on the CCD instead of altitude and wavelength. The second of these two changes has the greatest impact because of the interpolation of the kernels to the grid of the CCD. The interpolation can be sensitive to errors, especially when the kernel is strongly peaked at the tangent height, with large variations in amplitude as a function of altitude.

We note also that the PBR algorithm implementation includes the effects of CCD smear, which is modeled on an iterative basis. CCD smear occurs when charge is deliberately shifted from one capacitive bin to the next, in the process of reading out the CCD. Smear is most prominent in the short wavelength region (< 300 nm) of each image on the CCD. Consequently, its impact is greatest for the parts of the ozone profile that are derived from the UV channels, i.e., the high altitude portion of the ozone profile.

One modification that remains to be implemented is to construct UV ozone data vectors that are spectrally averaged to reduce the effects of low SNR. This process should also diminish the effects from tangent-height normalization. The lowest SNR pixels are located in the upper-right sections of each image in Fig. 8.These effects can further corrupt the UV ozone retrieval at high latitudes where ozone absorption is small.

3.2 RSAS Retrieval Results

The RSAS retrieval problem has been successfully implemented into the PBR algorithm, and tested with simulations including full stray light from the PSF convolution. Rather than using the instrument bore-sight altitude as the quantity to be perturbed in the retrieval, it was found via testing that a better quantity to use is the actual mapped altitude for the instrument (z_{ij}), where *i* and *j* are the CCD spectral and spatial indexes, respectively. Applying an induced altitude shift $\Delta z = \pm 1$ km, directly to z_{ij}, yields the correct change in signal of about 10-15% (from pure Rayleigh scattering). Figure 9 illustrates this point for one column (#390) from the low gain, central image, for pixels between altitudes of 35 and 45 km (pixels 272-284). The central wavelength of this column is ~350 nm. To isolate the effect of the applied Δz, these profiles do not include simulated instrument noise or stray light.

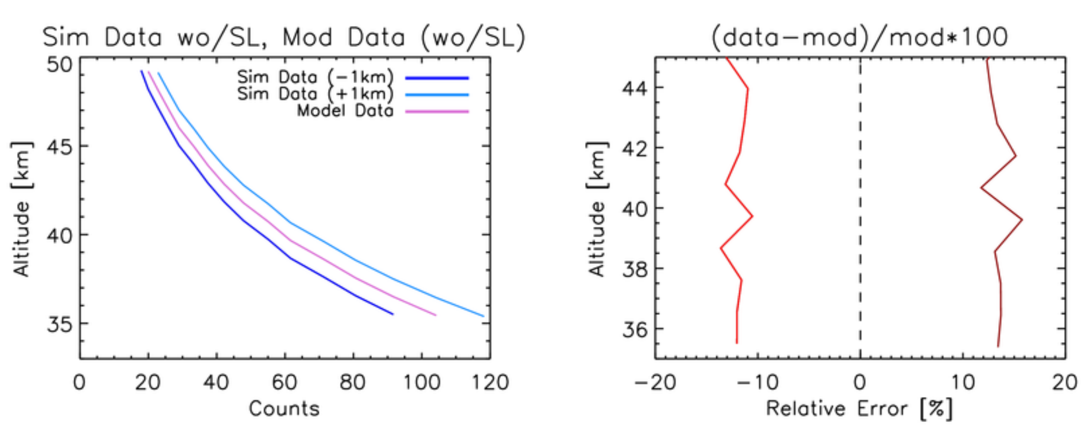

Figure 9 — Profiles of pixel column 390 (low-gain image). All profiles shown are without stray light and instrument noise. Left panel: Column-390 profiles from induced truth Δz. (Note: model data is Δz = 0.0 km.) Right panel: Percent differences (red line Δz = -1.0 km, and dark-red line Δz = 1.0 km).

Table 1 presents results of altitude-offset retrievals from the PBR algorithm using noisy simulated data with induced altitude errors of -1.0, 0.0, and 1.0 km. Inputs to the simulation included a mid-latitude ozone profile with a moderate springtime aerosol field, at a solar-zenith and a solar-azimuth angle of 50° and 40°, respectively. Each slit in the spectrometer sees the same atmospheric view. For each case a set of 30 random noise distributions were generated. The results shown in Table 1 indicate that a mean altitude error of ~ 100 m persists when stray light convolution is ignored in the retrieval. This error can be much larger, depending on the atmospheric conditions and viewing geometry. The bottom row of Table 1 shows that by including the full stray light modeling in the retrievals decreases the altitude errors to negligible levels, on the order of 10 m or less.

Test Case	$\Delta z = -1.0$ km	$\Delta z = 0.0$ km	$\Delta z = +1.0$ km
	Mean Error in Retrieved Offset (m)		
Without SL	121.05	102.41	113.68
With SL	10.97	2.76	10.66

Table 1 – RSAS retrievals utilizing three different *truth* offsets to create simulated data. Shown in the table are averaged mean errors (Row 1 excluded PSF convolution, and Row 2 included a full PSF convolution). The individual mean errors (in meters) were calculated the following way: $m = |\Delta z_{retrieval} - \Delta z_{truth}|$.

3.3 Surface Reflectance Retrieval Results

Fig. 10 shows results for the retrieval of surface reflectance, using one run from the 72-run ensemble with 50 different random-noise fields having been applied to output from that run. The data were generated with a mid-latitude ozone profile and a springtime aerosol field, at solar-zenith and solar-azimuth angles of 50° and 40°, respectively. As always, all of the spectrometer slits are seeing the same atmospheric scene. The *truth* surface reflectance was fixed at 0.15, while the first guess and *a priori* value were set to 0.30. All other retrieval inputs were held fixed at the values used for generating the simulated CCD maps. The top panel displays the retrieval results when stray light (convolution by the full PSFs) is included in the retrieval forward model. The statistical mean and sample standard deviation are shown in the upper right-hand corner. The bottom panel shows retrieval results when the PSF convolution is turned off in the retrieval forward model, corresponding to stray light being ignored in the retrieval even though it is present in the data. A significant change of ~38% is noted when stray light is not taken into account in the retrieval. These simulations were repeated using different values for the *a priori* and the same results were obtained, indicating that there is little or no *a priori* dependence in the surface reflectance retrieval.

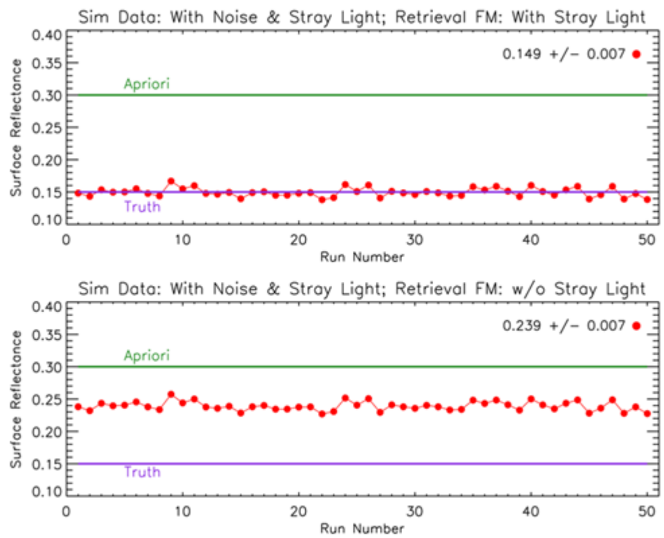

Figure 10 . Surface reflectance retrieval results using the pixel-based algorithm. The mean and standard deviation of the retrieval error are shown in the upper right hand corner of each plot. Top panel: Results obtained when the full point-spread function (PSF) convolution is included in the retrieval forward model. Bottom panel: Retrieval forward model does not include PSF convolution.

3.3 Chappuis-band Ozone Retrieval Results

Chappuis-band ozone retrievals were exercised extensively using the PBR algorithm. Figure 11 shows results from an initial analysis used to characterize the effect of instrument noise on the retrievals. The simulated data were generated using the same input parameters that were used in the RSAS and surface reflectance retrievals of Sections 3.1 and 3.2. Fifty files of noisy simulated data were produced from this single data set. Apart from the ozone *a priori* profile, which differed from the ozone *truth* profile, all retrieval parameters were the same as those used in generating the simulated data. The top panel displays results from a retrieval that does not include stray light in the retrieval forward model. The bottom panel includes stray light in the retrieval forward model. Between 10-20 km, the error profiles between both retrievals have the same morphology, but above 20 km the retrieval that included stray light more accurately captures the shape of the *truth* profile. Instrument noise does not adversely affect retrieval performance of Chappuis-band ozone.

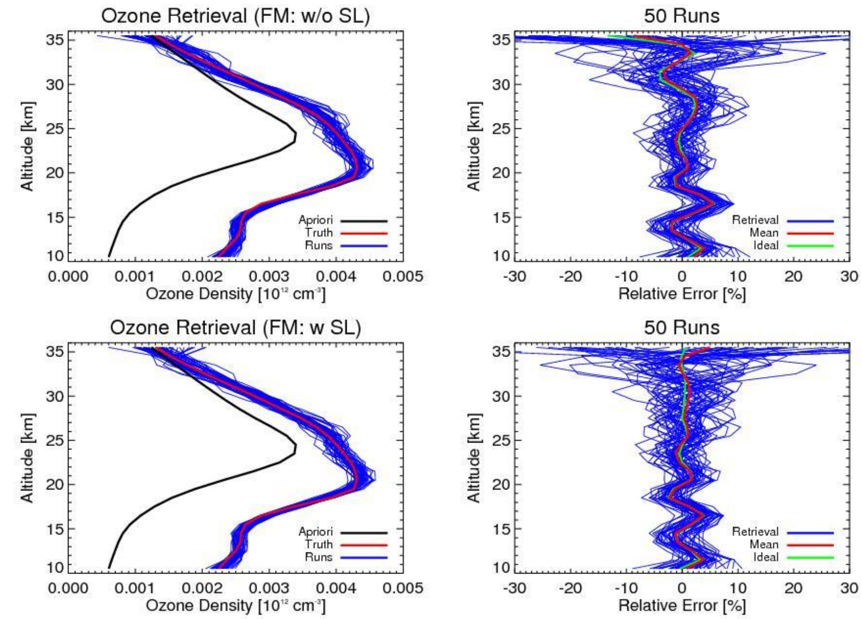

Figure 11 -- Chappuis-band ozone retrievals. Simulated data included instrument noise, CCD smear, and a full PSF convolution. In the 50 runs, a fixed set of atmospheric and viewing geometry inputs were used. These inputs were the same inputs that were used for the RSAS and surface reflectance retrievals in Sections 3.1 and 3.2. The green curve in the right-hand plots represents an error profile from an ozone retrieval where all input parameters were known.

Figure 12 shows results from a more stringent set of tests of Chappuis-band ozone retrieval performance. Here the standard 72-run data ensemble was used. Instrument noise was suppressed for these tests and error profile plots are shown only for the 150, 350, and 550-DU cases. The top panels display ozone retrieval results when all retrieval parameters equal their *truth* values, but with and without stray light included in the retrieval forward model. The results from retrievals that included stray light in the forward model (yellow line) are essentially error-free. However, when the stray light is turned off in the retrieval forward model, significantly larger errors appear. This bias becomes larger for lower-ozone cases, reaching 10-20% in the lower stratosphere for the low-ozone 150 DU case in the left panel. This is because less absorbing ozone means more light entering the instrument, and hence more stray light, together with weaker ozone absorption features, hence less signal to noise. This phenomenon has negative implications for retrieving high-latitude ozone, where ozone amounts are comparatively smaller.

The bottom set of plots corresponds to those in the top set, but for conditions closer to those that would occur in a realistic application of the algorithm. The first guess and *a priori* ozone profile used is different than the truth profile. A retrieved surface reflectance was used, but the input aerosol fields were the same as was used in creating the simulated data. The error-profile patterns for retrievals without stray light have the same morphology as those shown in the top row, although the magnitude of the errors is increased slightly.

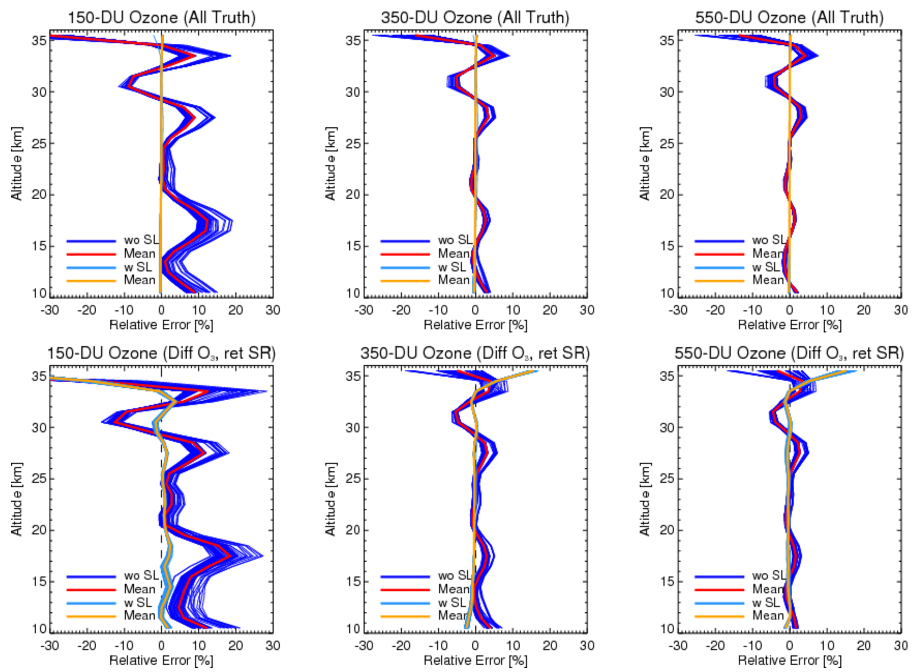

Figure 12 -- Retrieval results of visible ozone using the PBR algorithm. Top panels: All relevant atmospheric inputs were held at their *truth* values; Bottom panels: Relevant inputs included a different ozone *a priori*, a retrieved surface reflectance, a retrieved RSAS value, and the aerosol field was held at the *truth* aerosol field.

3.4 Hartley-Huggins Bands Ozone Retrieval Results

It might seem that retrieving ozone at higher altitudes using the LP UV channels would be easier than the low altitude Chappuis band retrievals. Aerosols are scarcer, reflections from the surface are weaker, and the use of on-off spectral pairs further reduces the influence of aerosols and scattering by the surface and clouds. But the instrument's SNR is smaller in this spectral region, and stray light contributes a larger fraction of the signal, so retrieving the ozone profile at these tangent altitudes is challenging.

Fig. 13 shows results analogous to Fig. 12, but for altitudes between 30 to 60 km the UV band ozone retrievals The first guess parameter values were the same as for Chappuis-band retrievals in Section 3.3, except that in the bottom row the aerosol first guess no longer coincides with the aerosol field used for simulating the data. The mean profiles of retrieval error shown in the top panels are reminiscent of those from the Chappuis-band retrievals displayed in Fig. 12. Only here the scatter is much greater and the bias introduced by ignoring stray light in the forward model is much larger than at the lower altitudes. As before, the same qualitative behavior is seen, with much larger errors occurring for low-ozone profiles. Even where all *truth* inputs are used, the retrieval does not perform well at the highest altitudes because there is so little ozone there.

In the bottom row of plots it is best to concentrate on the light blue curves, and the yellow mean curve, because it is already clear that ignoring stray light will produce unusable results. The mean profile and the scatter for the 150-DU true profile are significantly worse than in the earlier plots for that true profile. Using a retrieved altitude offset and surface reflectance and aerosol worsens the effect of the especially small SNR for this profile. As the amount of ozone increases (350 and 550 DU), the retrieval performs better, except at the highest altitudes.

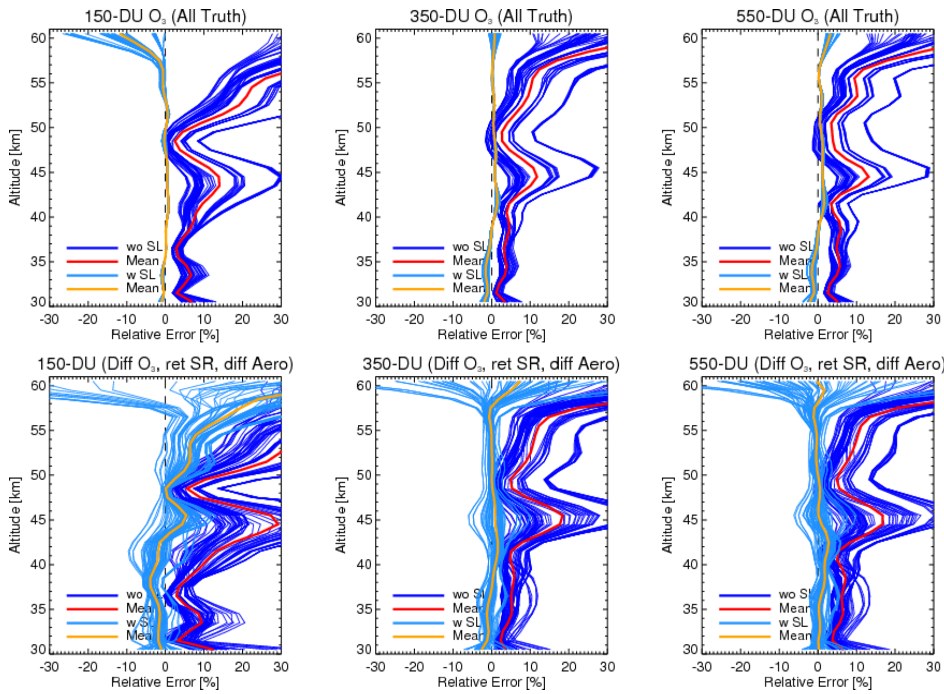

Figure 13 – UV ozone retrievals for all *truth* inputs (top panels), and a combination of retrieved and unknown inputs (bottom panels).

CONCLUSIONS

An alternative retrieval algorithm based on a forward method for correcting stray light has been developed for the OMPS/LP instrument. This method is implemented by using a two-stage forward model, where the first stage is a solar limb-scatter RT model that includes spherical geometry, all orders of scattering, and polarization. The second stage is an instrument model that maps the radiance field produced by the RT code through the spectrometer to create a realistic CCD signal map. All instrument effects, including a realistic stray light simulation, are realistically simulated within this combined forward model, resulting in modeled signals that can be compared directly with the uncalibrated data obtained directly from the downlinked CCD pixels.

All retrieval steps except for the aerosol component have reached a mature stage and are functional in this alternative algorithm. The aerosol retrieval is working and was being integrated into the PBR algorithm at the time work stopped on the project. Simulations indicate that for moderate to high ozone levels the UV and visible ozone retrievals will work well as currently implemented in the algorithm. However, low-ozone conditions will present problems at several levels for both UV and visible retrievals, with the UV problem being the most severe due to low ozone, low pixel signal values, and increased stray light contamination.

The alternative approach to stray light mitigation outlined here may offer the best prospect for addressing a problem that plagues many remote sensing instruments. With the OMPS/LP, stray light is a severe problem. Previous attempts to correct this problem at the SDR level proved to be problematic, mostly because of errors and instability introduced by signal deconvolution. With the application of the PBR algorithm, retrievals of geophysical constituents can proceed in a mathematically sound way to obtain stable and accurate ozone EDRs.

REFERENCES

1. R.J. Thomas, C. A. Barth, D. W. Rusch, and R. W. Sanders, "*Solar Mesospheric Explorer Near-Infrared Spectrometer: Measurements of1.27-μm Radiances and the Inference of Mesospheric Ozone*", J. Geophys. Res., 89, 9,569-9,580 (1984). D. W. Rusch, G. H. Mount, C. A. Barth, R. J. Thomas, and M. T. Callan, "*Solar Mesospheric Explorer Ultraviolet Spectrometer: Measurements of Ozone in the 1.0-0.1 mb Region*," J. Geophys. Res., 89, 11,677-11,687 (1984).
2. D E. Flittner, P. K. Bhartia, B. M. Herman, *O_3 Profiles Retrieved From Limb Scatter Measurements: Theory*, *Geophys. Res. Lett.*, 27, 17, doi:10.1029/1999GL011343 (2000). R.D. McPeters, S. J. Janz, E. Hilsenrath, T. L. Brown, D. E. Flittner, and D. F. Heath, "The retrieval of O3 profiles from limb scatter measurements: Results from the Shuttle Ozone Limb Sounding Experiement", *Geophys. Res. Lett.*, 27, 17, doi:10.1029/1999GL01132 (2000).
3. E. J. Llewellyn, *et al*., "*The OSIRIS Instrument on the Odin Spacecraft*," Can. J. Phys., 82, 411-422 (2004).
4. C. von Savigny et al. ,"Detection and mapping of polar stratospheric clouds using limb scattering observations", Atmos. Chem. Phys. Discuss., 5, 7169-7190, (2005).
5. D. F. Rault, "Ozone profile retrieval from Stratospheric Aerosol and Gas Experiment (SAGE III) limb scatter measurements", J. Geophys. Res., 110, D09309, doi:10.1029/2004JD004970.
6. Versions of the ATBD for the OMPS Limb Profiler prior to about 2004. Contact the JPSS office about availability.
7. J. W. Bergman, L. E. Flynn, J. S. Hornstein, J. D. Lumpe, "Using regression models to enhance signals in a scattered radiation field: Reducing stray light contamination in the limb profiles of the Ozone Mapper Profiler Suite (OMPS), AGU Meeting, 11-15 December 2006. John W. Bergman, Didier F. Rault, John S. Hornstein, Jerry D. Lumpe, "Reducing stray light contamination in the limb profiles of the Ozone Mapper Profiler Suite (OMPS)", American Meteorological Society Meeting, 20-24 January 2008. J. S. Hornstein, T. D. Eden, J. Lumpe, and D. F. Rault, "An Alternative EDR Retrieval Strategy for the NPP OMPS Limb Profiler," *American Geophysical Union, Fall Meeting,* 2009.
8. C. D. Rodgers, *Inverse Methods for Atmospheric Sounding, Theory and Practice, Vol. 2,* World Scientific, Hackensack, N.J. 07601 (2008).
9. J. D. Lumpe, R. M. Bevilacqua, K. W. Hoppel, And C. E. Randall, *POAM III Algorithm and Error Analysis*, J. Geophys. Res. 107(D21), 4575, 10.1029/2002JD002137.
10. J. D. Lumpe, L. E. Floyd, L. C. Herring, S. T. Gibson, and B. R. Lewis, *Measurements of Thermospheric Molecular Oxygen From the Solar Ultraviolet Spectral Irradiance Monitor*, J. Geophys. Res., 112, D16308, doi:10.1029/2006JD008076.
11. M. H. Bettenhausen, *A Nonlinear Optimization for WindSat Wind Vector Retrievals,* IEEE Trans. On Geosci. And Remote Sensing, vol. 44, no. 3, pp. 597-610 (2006).
12. B. Herman, T. Caudill, D. Flittner, K. Thome, and A. Ben-David, "Comparison of the Gauss-Seidel spherical polarized radiative transfer code with other radiative transfer codes", *Appl. Opt*., 34, 4563-4572m (1995).

Appendix A – OMPS/LP instrument simulator

The instrument simulator for the OMPS limb-profiler (LP) was developed as part of an end-to-end testing platform for OMPS-LP algorithm development. The simulator provides a high fidelity, comprehensive description of all fundamental optical and electronic components of the OMPS-LP instrument, allowing for the simulation of realistic LP level 0 data (CCD pixel signals) given an input atmospheric radiance field. This appendix provides a brief overview of the features of the model; a much more detailed description of the model including derivations of the algorithms, source documentation for instrument specifications and comprehensive instructions for compiling and running the FORTRAN source code, are given in [1].

The instrument simulator calculates measured electron counts for each pixel on two CCD arrays (long and short integration times), each containing six images (three slits, two difference gains each) using as input three sets (one for each slit) of radiance data (radiant power per area per solid [viewing] angle per wavelength increment; W m^{-2} sr^{-1} nm^{-1}) as a function of wavelength and tangent altitude as seen from the satellite for a given viewing geometry. There are four fundamental components of the model.

1. The optics convert the three sets of input radiance data into six images (high and low gain for each slit) of radiant power (W) projected onto the focal plane. The effective area of the instrument's low gain aperture determines the radiant power density (W sr^{-1} nm^{-1}) and multiplication by the high-to-low gain ratio determines the radiant power density for the high gain aperture for each slit. The optical throughput efficiency determines how each of these images is attenuated and the point spread function determines how light from each image is scattered as the light passes through the optics. The projection of the slit onto the focal plane then determines the radiant power distribution onto the focal plane.
2. The conversion of radiant power (W) into photon count rates (photons per unit time; s^{-1}) entering each pixel on the CCD includes: dividing the radiant power by the energy per photon, multiplying by the long and short integration times (obtaining two CCD arrays of photon counts), adding shot noise that arises from the fact that photon emission from a radiating source is random – following Poisson statistics, and adding smear: photons that are accumulated during the transfer of electrons to the non-illuminated part of the focal plane. Smear also has a random component due to the shot noise for those photons.
3. The imager governs the conversion of photon counts entering each pixel into electrons stored in each pixel of the two CCD arrays. Photon counts are attenuated by the quantum efficiency (due to imperfect photo-electric effect) of the pixels. Then, we add dark current (which has both systematic and random components), readout noise, and clocking noise.
4. The measurement of the number of electrons stored in each pixel is subjected to electronics effects. There is a systematic component (the electronics 'offset', specified via charge transfer efficiency) that has a nonlinear dependence on the number of stored electrons and random components due mostly to amplifiers and the analog-to-digital conversion.

Components of the model

Here, we present an outline of the various instrument features in the order that they are calculated in the simulator. Details for each calculation are given in [A1].

1. **Initialization:** Read input parameters: logical switches, input and output file names. Read input data: 3 radiance fields, one for each slit, and set up auxiliary arrays.
2. **Optics:** Converts radiance entering the instrument into the average rate at which photons enter each pixel.

   a. **Effective aperture**: Used as a tuning parameter for now because we do not have sufficient information regarding the radiometric calibration of the instrument. This parameter accounts for components of the optics that impact the radiant energy uniformly at all wavelengths. It converts radiance (W m^{-2} sr^{-1} nm^{-1}) to radiant power density (W sr^{-1} nm^{-1}).
   b. **Create high gain image**: Create high gain images for each slit by multiplying each image by a constant. There are now 6 images.
   c. **Apply optical through put:** Accounts for wavelength-dependant attenuation of the signal (absorption of energy in the optics).
   d. **Apply nominal-to-measured throughput correction:** Values of throughput are from a theoretical calculation. This correction adjusts the throughput to match measured performance of the instrument.
   e. **Quantum efficiency:** Actually an imager effect, but to include its wavelength dependence, it is calculated here. Once the point-spread-functions (PSFs) are applied, wavelength information for individual photons is lost and, so, calculating quantum efficiency during the imager calculations is problematic.
   f. **Radiance to photon conversion:** Convert radiant power density to photon density rate (photons sr^{-1} nm^{-1} s^{-1}).
   g. **Project images onto CCD:** Integrate photon density rate over wavelength and solid angle and project the signal onto the CCD. Converts photon density rate to the photon count rate (photons s^{-1}) for each pixel.
   h. **Apply stray light filters:** Create pre-filtered data that will act as the 'photon sources' (see detailed description for application of the PSF for a clarification of this term) for PSF applications.
   i. **Apply point spread functions (PSF):** Accounts for internal scattering in the instrument. Uses pre-calculated PSFs and a method that utilizes coarse-grid PSFs in the (low-amplitude) tails to reduce computational cost.
3. **Imager:** Converts the rate at which photons enter each pixel into the number of electrons stored in each pixel of the 2 CCD arrays.
   a. **Multiply photon count rate by integration time:** Converts photon count rate to photon counts for two CCD arrays of photon counts (short and long integration)
   b. **Apply smear**
   c. **Add photon shot noise (includes smear noise)**
   d. **Add dark current bias**
   e. **Add dark current, read, and clocking noise**
   f. **Saturate the pixels:**
4. **Electronics:** Converts the number of electrons stored in each pixel to the number of electrons measured for each pixel.
   a. **Electronics offset:** Charge transfer efficiency**;** Contains nonlinearity
   b. **Electronics noise:** Amplifier and analog-to-digital conversion noise.

A1. J. W. Bergman and D. F. Rault, "*OMPS-LP Instrument Simulator Version 2*", CPI internal technical report (2007).

GLOSSARY

BATC	Ball Aerospace and Technology Corporation
CPI	Computational Physics Inc.
EDR	Environmental Data Record
IPO	Integrated Program Office
IR	Infra-Red (0.76 µm to 1 mm wavelengths[1])
LP	Limb Profiler
NASA	National Aeronautics & Space Administration
NIR	Near-Infrared
NPOESS	National Polar-orbiting Operational Environmental Satellite System
NPP	NPOESS Preparatory Project
NRL	Naval Research Laboratory
OSIRIS	Optical Spectrograph and Infrared Imaging System
PBR	Pixel-Based retrieval
POAM	Polar Ozone and Aerosol Measurement
PSF	Point Spread Function
RSAS	Rayleigh Scattering Altitude Registration
RT	Radiative Transfer
SBUV	Solar Backscatter Ultra-violet
SDR	Sensor Data Record
TOMS	Total Ozone Mapping Spectrometer
UV	Ultraviolet (100 nm to 400 nm wavelengths[2])
VIS	Visible (380 nm to 760 nm wavelengths[3])

[1] Summary of ISO 21348 wavelengths and compliance criteria, http://www.spacewx.com/pdf/ SET_21348_2004.pdf
[2] Summary of ISO 21348 wavelengths and compliance criteria, http://www.spacewx.com/pdf/ SET_21348_2004.pdf
[3] Summary of ISO 21348 wavelengths and compliance criteria, http://www.spacewx.com/pdf/ SET_21348_2004.pdf